Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 225 (2023) 1928–1936

Procedia Computer Science

www.elsevier.com/locate/procedia

27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

A Study of User Minds for Mobile Payment using Text Mining

Yasunobu Kino[a,*], Caroline S.L. Tan[a], Yi Zhu[b]

[a]*University of Tsukuba, 3-29-1 Otsuka, Bunkyo-ku, Tokyo, 112-0012, Japan*
[b]*Lancaster University, Lancaster, LA1 4YW, United Kingdom*

Abstract

Recently, the usage of mobile payments in Japan has increased. Some users have expressed positive views about this technology, while others hold negative views. This study aimed to examine the attitudes of both users and non-users toward mobile payments by applying text mining on Japanese comments posted on Twitter. The data were collected twice in February 2023. Through the analysis, the study identified keywords related to positive, negative, and risk-related sentiments to gain insight into the expressed opinions. The findings revealed that positive statements highlighted the convenience of mobile payments, ease of ordering, and the value offered through loyalty points. Negative themes encompassed issues like settlement problems and wireless connection troubles. Regarding risks, there were statements such as personal data leaks and various security issues. Furthermore, a comparison between mobile payments and credit cards indicated that mobile payments received fewer risk-related statements and more positive statements compared to credit cards.

Keywords: Cashless payment; Text Analysis; Customer Experience

* Corresponding author.
E-mail address: kino@mbaib.gsbs.tsukuba.ac.jp

1877-0509
10.1016/j.procs.2023.10.183

1. Introduction

In recent years, the adoption of mobile payment in Japan has seen a rapid increase. Consumers are constantly seeking and comparing different mobile payment methods to determine their preferred choice. This study aims to analyze the content and sentiments expressed by individuals on Twitter regarding mobile payments.

The world is currently moving toward a cashless society, and in Japan, a range of mobile payment options, including PayPay and Suica, are gaining traction in the market. While mobile payments offer convenience, they also present challenges such as personal information leakage, identity theft, loss of cell phones, and data extraction, which contribute to a slower adoption rate of mobile payments in Japan. In light of this situation, this study analyzes Twitter posts to gain insights into the opinions of mobile payment users.

The analysis conducted in this study focuses on the Japanese language. Unlike English or Spanish, Japanese is predominantly used within Japan, except for Japanese individuals living abroad. Additionally, although Japanese and Korean languages are considered grammatically similar, they have distinct differences that can be easily recognized when looking at the written characters. Therefore, it is reasonable to assume that most of the content spoken in Japan is the same as that in Japan. Given these characteristics of the Japanese language, this study analyzes Japanese speech as a representative of Japanese opinions. With mobile payments gaining popularity in Japan, this study employs text mining techniques to analyze comments posted in Japanese on Twitter, providing insights into the perspectives of both users and non-users.

2. Analytical process

To identify the characteristics of the tweets posted by mobile payment users, we performed several operations, as shown in Figure 1. The data were collected from Twitter on February 15 and February 27, 2023 using specific keywords such as "PayPay," "d payment," "LINE Pay," "Osaifu-Keitai," "QR code payment," "cashless payment," "mobile payment," and "credit card." The first four keywords ("PayPay," "d payment," "LINE Pay," and "Osaifu-Keitai") represent commonly used mobile payment services in Japan. We obtained a maximum of 1,000 tweets for each keyword on the specified dates, resulting in a total of 14,519 collected tweets.

The collected tweets included not only personal tweets but also corporate tweets from mobile phone companies or shops; these corporate tweets were considered to be advertisements. Therefore, we excluded tweets that met the following four conditions: (1) Tweets with repeated information (likely to be posted by companies), (2) tweets posted by official accounts of companies (self-declared to be posted by companies). (3) Promotional tweets posted by companies with phrases such as "excuse me from search." (4) Tweets not written in Japanese, as they fell outside the scope of our study.

In cases where tweet evaluation was challenging, we carefully read the content and removed those tweets that appeared to be posted by companies rather than individuals. Consequently, a total of 7,955 tweets remained and were used for the analysis.

Although the original data of this study are in Japanese, our paper is written in English. Therefore, we performed text mining using two approaches: one for Japanese analysis and the other for English analysis. The data for the English analysis were automatically translated from the original Japanese data using Google Translate. We confirmed that there were no significant differences in the results derived from the Japanese and English data, as the tweet sentences were divided into keywords.

We used KH Coder in our analysis [1]. In addition, we employed ChaSen from the Nara Institute of Science and Technology as a morphological analysis tool for the Japanese text and Stanford POS Tagger for the English text. Our study builds upon previous studies [2–6] that used text analysis techniques.

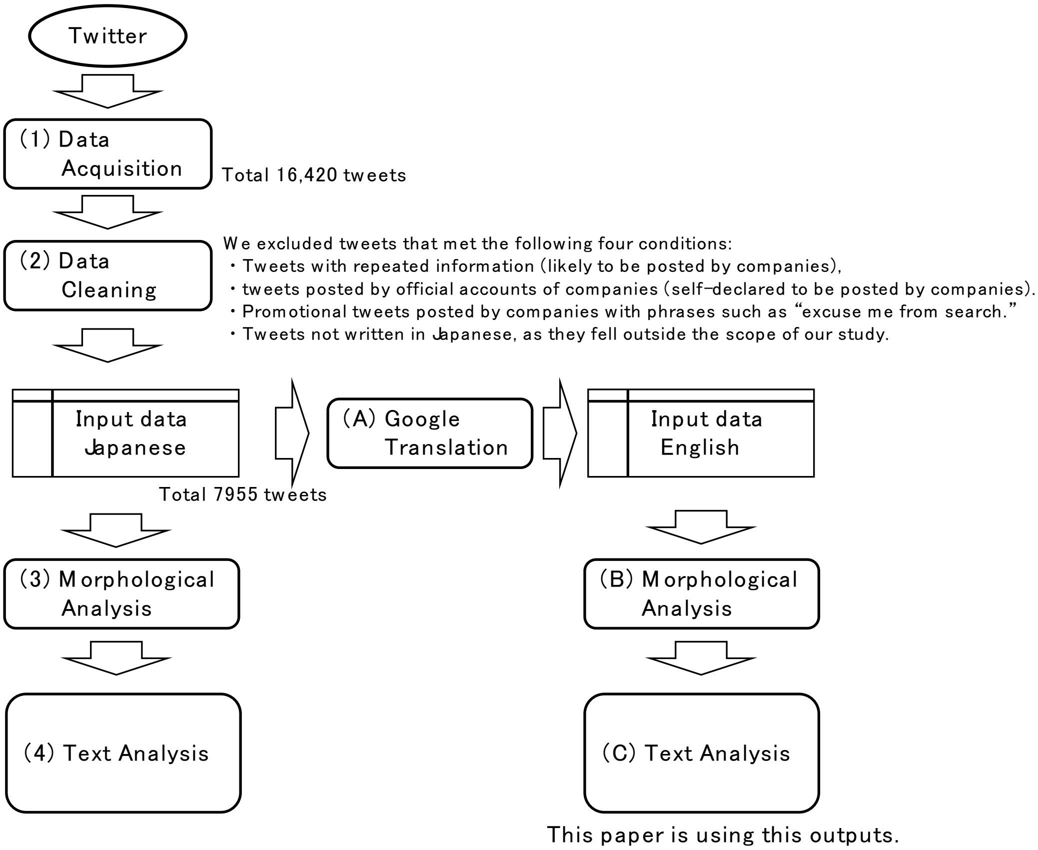

Fig. 1. Analytical Process

3. Results of Analysis

3.1. Frequent Keywords

Morphological analysis was performed on the cleaned English data using the Stanford POS Tagger from KH Coder. The analysis involved splitting words such as "credit card" into two separate words, "credit" and "card." However, treating them as a single word proved more advantageous for the analysis. Such words were included in the Force Extract words list, as indicated in the TAG column of Table 1. Additionally, words such as "be," "do," and "have," which are used in different contexts in English were added to the Force Ignore words list because they interfere with the analysis.

Table 1 lists the top 25 nouns, adjectives, adverbs, verbs, and TAGs. Note that the term "point" holds the third position among the nouns and refers to the concept of "reward points."

Table 1. Frequent keywords for each part of speech

Noun		TAG		Adjective		Adverb		Verb	
payment	5221	credit_card	2099	cashless	1551	not	3852	use	2689
card	1239	QR_code	1235	mobile	693	so	794	pay	2278
point	1239			good	396	also	673	think	831
cash	886			able	353	only	647	make	815
yen	831			first	326	even	559	get	797
store	792			possible	298	now	357	go	788
smartphone	592			other	272	too	350	want	670
money	589			such	264	just	334	buy	576
pay	577			electronic	254	really	292	seem	421
people	426			many	245	well	161	know	416
wallet	383			convenient	235	still	153	please	363
number	363			more	225	more	149	say	352
day	348			various	166	here	144	take	350
ticket	334			new	159	very	143	try	290
fee	331			same	159	finally	123	come	285
app	321			last	152	back	122	charge	281
method	299			free	148	recently	122	see	276
today	273			high	134	as	121	look	264
service	264			next	134	again	114	wonder	261
order	234			available	131	however	105	support	258
company	233			much	125	much	103	start	250
account	230			little	117	already	99	become	236
campaign	226			easy	113	right	98	feel	235
amount	220			long	112	always	95	like	222
information	220			troublesome	112	never	86	receive	222

3.2. Co-occurrence Network

Following the extraction of frequent keywords, a co-occurrence network diagram was generated (Figure 2 shows a diagram using all the data). This diagram visually represents the connection between keywords that have a strong tendency to occur together in sentences. Keywords linked by lines indicate a strong association in terms of their descriptions. By analyzing groups of keywords connected by lines, potential themes were derived.

In Figure 2, the key themes of this article, namely “payment,” “cashless,” and “use,” are depicted in the center. Noteworthy characteristic keywords include “point” and “get” due to the presence of rewards and competition among different mobile payment services, leading users to anticipate earning points.

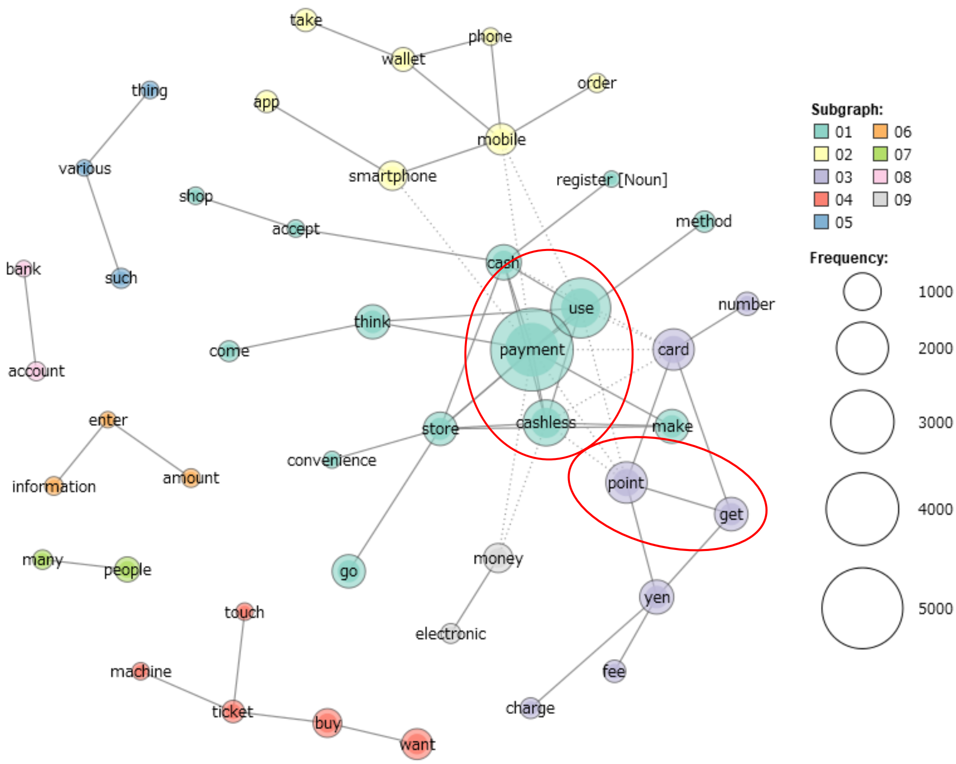

Fig. 2. Co-occurrence Network Diagram

Subsequently, by examining all the frequent keywords (full version of Table 1), keywords associated with "positive," "negative," and "risk" were identified. With further analysis of the actual usage of these keywords, those effectively expressing "positive," "negative," and "risk" sentiments were sorted. During the sorting process, keywords with dual meanings encompassing both "positive" and "negative" connotations were excluded. For example, the word "good" was omitted because it often appears in negative contexts as "not good" and therefore cannot be considered to have a truly "positive" connotation. The summarized results of this aggregation are as follows.

Keywords related to Positive: convenient, easy, fine
Keywords related to Negative: trouble, troublesome, bad, useless, difficult
Keywords related to Risk: risk, risky, leak, security, failure, fail

KH Coder aggregates multiple keywords and analyzes them based on their shared meaning. Utilizing the "positive," "negative," and "risk" keywords identified in this study, co-occurrence network diagrams were created using keywords related to the above-mentioned keywords (often used together in sentences), as illustrated in Figures 3, 4, and 5.

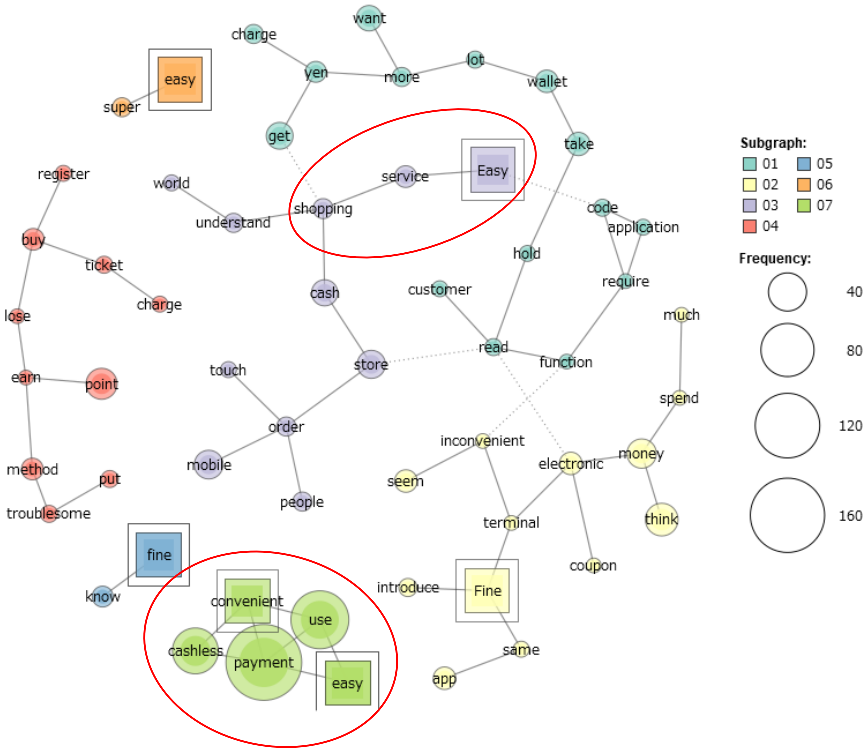

Fig. 3. Co-occurrence Network Diagram (Positive)

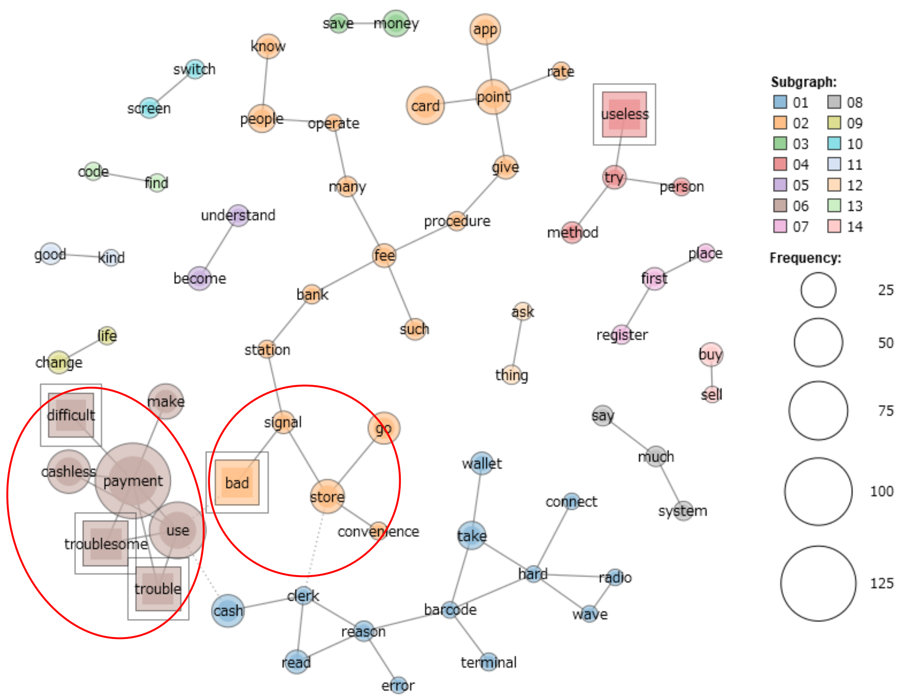

Fig. 4. Co-occurrence Network Diagram (Negative)

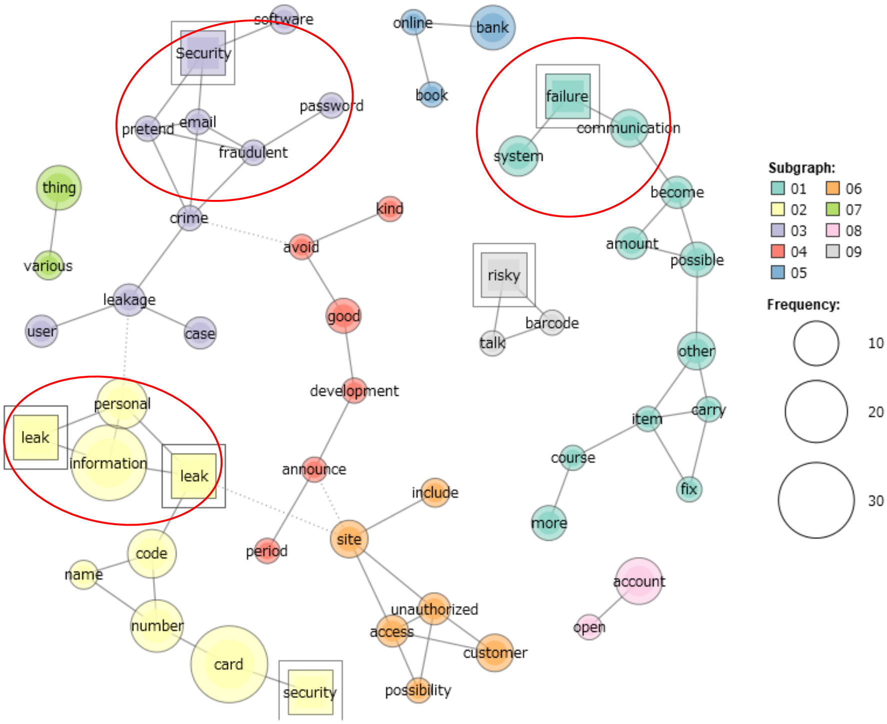

Fig. 5. Co-occurrence Network Diagram (Risk)

In Figure 3 (Positive), a cluster comprising “convenient,” “easy,” “cashless,” “payment,” and “use” can be observed, with frequent mentions of the comment “using cashless payment is easy and convenient.” There is another cluster that consists of “easy,” “service,” and “shopping,” where “service” is frequently associated with “payment” as “payment service” and “shopping” is linked with “payment service” as “easy.”

In Figure 4 (Negative), there is a cluster that consists of “payment,” “use,” “trouble,” “troublesome,” “difficult,” and “cashless,” describing difficulties and problems encountered when using cashless payments. There is also a cluster of “bad,” “signal,” “shop,” and “convenience.” When we checked the original text, we found statements such as "I couldn't use it at the convenience store because of a bad signal."

In Figure 5 (Risk), a cluster of keywords consisting of “information,” “leak,” and “personal” can be observed. This cluster refers to “leaks of personal information.” Notably, within this cluster, there are two instances of the word "leak," both as a noun and an adverb. Furthermore, there are keywords such as "security,” “pretend,” “email,” “fraudulent,” “password,” and “software” indicating the different security issues being discussed. The significant occurrence of keywords such as "failure," "system," and "communication" indicate the prevalence of various problems associated with failures, malfunctions, and communication-related challenges.

3.3. Relationship between Extracted Word and Positive/Negative/Risk

To examine the relationship between the words used to extract the tweets (words related to mobile payments) and the number of sentences related to Positive/Negative/Risk, a cross-tabulation table (Table 2) was created.

Figure 6 illustrates the relationship between each keyword (horizontally) and the "positive," "negative," and "risk" keywords used to retrieve data from the tweets. The frequency of occurrence for each word is presented both horizontally and vertically in relative terms. While each specific service (Services A–D) possesses distinct characteristics and exhibits varying occurrence patterns, our focus here is on the relationship between credit cards and mobile payments. The trend in the occurrence of risk-related keywords suggests that risk is mentioned more frequently in relation to credit cards compared to mobile payments. Additionally, the trend in the occurrence of positive keywords is more prevalent for mobile payments than for credit cards. Although these data alone do not provide definitive conclusions, the inclination for mobile payment to have a more positive sentiment than credit card payment is a promising indication for the future adoption of mobile payment.

	Positive	Negative	Risk	sentences
Credit card	19 (0.93%)	30 (1.46%)	32 (1.56%)	2049
Payment Word	329 (1.90%)	265 (1.53%)	159 (0.92%)	17301
Service A	31 (1.58%)	25 (1.27%)	5 (0.25%)	1961
Service B	5 (1.29%)	5 (1.29%)	1 (0.26%)	389
Service C	20 (1.66%)	26 (2.16%)	10 (0.83%)	1206
Service D	15 (0.86%)	12 (0.69%)	4 (0.23%)	1742
Summary	419 (1.70%)	363 (1.47%)	211 (0.86%)	24648
chi-square value	19.447**	12.294*	30.888**	

Table 2. Number of sentences by each extracted word

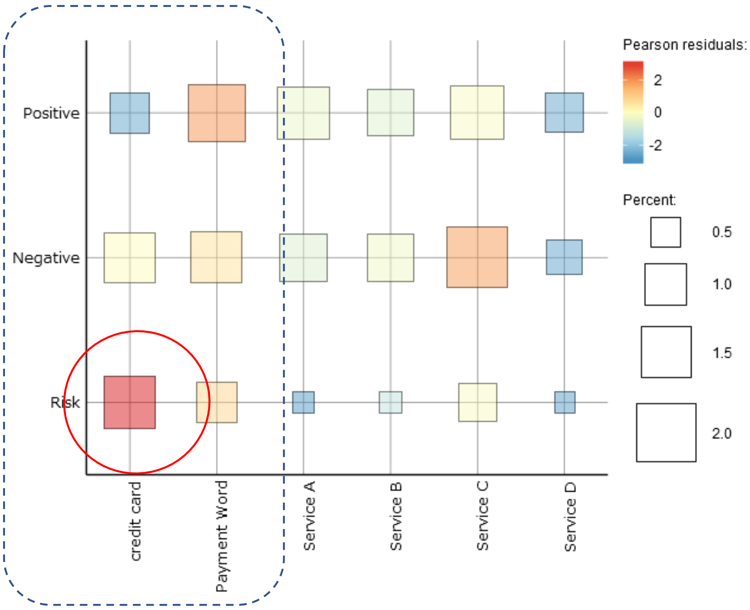

Fig. 6. Relationship between extracted words and Positive/Negative/Risk

3.4. Special Word

Japan is an earthquake-prone country, presenting a challenge that arises when using mobile payments in the absence of electricity. In the analyzed data, the keyword "earthquake" appeared 13 times. However, upon reviewing the original text, we found that most of the references to "earthquake" were related to the ability to use mobile payments to make donations toward the Turkish earthquake, which had occurred prior to data collection.

4. Conclusion

This study investigated attitudes toward mobile payments by performing the text mining of Japanese comments posted on Twitter. The data were collected twice, on February 15 and February 27, 2023. The selected keywords for data collection were “PayPay,” “d payment,” “LINE Pay,” “Osaifu-Keitai,” “QR code payment,” “cashless payment,” “mobile payment,” and “credit card.”

The collected tweets encompassed not only individual posts but also numerous posts that appeared to be advertisements from companies, such as mobile phone and advertising companies. Therefore, corporate and commercial tweets, as well as tweets in foreign languages, were excluded, resulting in 7,955 tweets available for analysis. The original data used in for this study were in Japanese; however, because our paper is written in English, we employed two systems for text mining: one for Japanese analysis and the other for English analysis. For the English analysis, the original Japanese data were translated using Google Translate. The analysis involved keyword segmentation, and it was confirmed that there were no significant differences between the Japanese and English analysis results. The analysis focused on identifying positive, negative, and risk-related keywords, as well as examining the sentiments expressed. Positive statements included statements such as “easy to settle,” “easy to order,” and “good value for money with points.” Negative themes included statements such as “settlement problems” and “communication problems.” Regarding risks, there were mentions of “personal data leaks” and “various security issues.” A comparison between credit cards and mobile payments indicated that mobile payments have fewer risk-related statements and more positive statements in comparison to credit cards.

While the data were collected for four specific services to address the research question, the differences in trends between these services were not analyzed. Moreover, as the data were only collected twice in February, it is possible that different trends may emerge if the data were collected over a more extended period. We therefore plan to conduct further research to delve into these issues.

References

[1] Koichi Higuchi (2014) "Syakai Cyosa no tameno Keiryo Text Bunseki." Nakanishiya syuppan.

[2] Yasunobu Kino (2016) "Skills for Global Human Resources and Structure Recognized by High School Students." Oukan 10 (2): 116–123.

[3] Daisuke Toyoshima, Yasunobu Kino, et al. (2016) "Extraction and Trend Analysis of the Word "Risk" in Journals and Proceedings." Journal of the Society of Project Management 18 (4): 20–25.

[4] Kiyoshi Sakamori, Kiyomi Miyoshi (2016) "Outlook of the program design for Project Management Professional School: Based on Text Mining Application." Bulletin of Advanced Institute of Industrial Technology (10): 79–84.

[5] Hiromi Asano, Koji Tanaka, Yoshikatsu Fujita, Kazuhiko Tsuda (2015) "Study on hiring decision: analyzing rejected applications by mining individual job placement data of Public Employment Security Offices." Procedia Computer Science 60: 1156–1163.

[6] Yasunobu Kino, Hiroshi Kuroki, Tomomi Machida, Norio Furuya, Kanako Takano (2017) "Text Analysis for Job Matching Quality Improvement." Procedia Computer Science 112: 1523–1530.